Title (en)

MULTI-MODAL SPEECH LOCALIZATION

Title (de)

MULTIMODALE SPRACHLOKALISIERUNG

Title (fr)

LOCALISATION DE PAROLE MULTIMODALE

Publication

EP 3791391 A1 20210317 (EN)

Application

EP 19723577 A 20190430

Priority

- US 201862668198 P 20180507
- US 201816020934 A 20180627
- US 2019029769 W 20190430

Abstract (en)

[origin: US2019341054A1] Multi-modal speech localization is achieved using image data captured by one or more cameras, and audio data captured by a microphone array. Audio data captured by each microphone of the array is transformed to obtain a frequency domain representation that is discretized in a plurality of frequency intervals. Image data captured by each camera is used to determine a positioning of each human face. Input data is provided to a previously-trained, audio source localization classifier, including: the frequency domain representation of the audio data captured by each microphone, and the positioning of each human face captured by each camera in which the positioning of each human face represents a candidate audio source. An identified audio source is indicated by the classifier based on the input data that is estimated to be the human face from which the audio data originated.

IPC 8 full level

G10L 17/10 (2013.01); **G06V 10/764** (2022.01); **G10L 17/00** (2013.01); **H04N 7/15** (2006.01); **H04N 23/90** (2023.01)

CPC (source: EP US)

G01S 3/801 (2013.01 - EP); **G01S 3/802** (2013.01 - EP); **G06F 18/2413** (2023.01 - EP); **G06T 7/70** (2017.01 - US); **G06V 10/454** (2022.01 - EP US); **G06V 10/764** (2022.01 - EP US); **G06V 40/165** (2022.01 - EP US); **G06V 40/166** (2022.01 - EP US); **G10L 17/00** (2013.01 - EP US); **G10L 17/10** (2013.01 - EP US); **G10L 21/0232** (2013.01 - US); **G10L 25/84** (2013.01 - US); **H04N 7/147** (2013.01 - EP); **H04N 7/15** (2013.01 - EP); **H04N 7/155** (2013.01 - EP); **H04N 23/611** (2023.01 - EP); **H04N 23/90** (2023.01 - EP); G01S 5/18 (2013.01 - EP); G06T 2207/30201 (2013.01 - US); G06V 40/172 (2022.01 - EP US); G10L 17/18 (2013.01 - US); G10L 2021/02166 (2013.01 - EP); H04N 23/90 (2023.01 - US); H04R 1/406 (2013.01 - EP US); H04R 2201/401 (2013.01 - EP); H04R 2430/20 (2013.01 - EP)

Designated contracting state (EPC)

AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

Designated extension state (EPC)

BA ME

DOCDB simple family (publication)

US 10847162 B2 20201124; **US 2019341054 A1 20191107**; CN 112088315 A 20201215; CN 112088315 B 20240402; EP 3791391 A1 20210317; WO 2019217135 A1 20191114

DOCDB simple family (application)

US 201816020934 A 20180627; CN 201980030506 A 20190430; EP 19723577 A 20190430; US 2019029769 W 20190430